An Improved Model for Live Migration in Data Centre Simulators

Vincenzo De Maio*, Gabor Kecskemeti†, Radu Prodan*
*Institute of Computer Science, University of Innsbruck
†MTA SZTAKI, Laboratory of Parallel and Distributed Systems
Email: {vincenzo,radu}@dps.uibk.ac.at, kecskemeti.gabor@sztaki.mta.hu

***Abstract*—Due to the difficulty of employing real data centres' infrastructure for assessing the effectiveness of energy-aware algorithms, many researchers resort to use simulation tools. These tools require precise and detailed models for virtualized data centres in order to deliver accurate results. In recent years, many models have been proposed, but most of them either do not consider energy consumption related to virtual machine (VM) migration or do not consider some of the energy impacting components (e.g. CPU, network, storage). In this paper, we propose a new model for data centre energy consumption that takes into account the previously omitted components and provides more accurate energy consumption predictions compared to other state-of-the-art solutions for paravirtualized VMs. We evaluated our model's accuracy in a comprehensive set of scenarios implemented in the DISSECT-CF simulator. Our analysis revealed a significant up to 42.5% improvement in accuracy for modelling data centre energy consumption compared to a similar state-of-the-art simulator.**

I. Introduction

Recently, Cloud computing has emerged as a computing paradigm by which computational power is hosted in data centres of specialised providers and rented on-demand of the users based on their occasional needs. Since power consumption has an increased significance for Cloud providers, they are more interested in optimising their data centre's energy efficiency to maximise their profit. One way of improving data centre energy efficiency is to maximise the utilisation of the *physical machines*(*PMs*), that are often under-utilised according to the study in [1]. For this purpose, data centre operators often apply a technique called workload consolidation that increases the resource utilisation by mapping computational tasks on a subset of the data centre's PMs and shutting down the rest (i.e., putting them in a low power state). Nowadays, computation is mostly running on *virtual machines (VMs)*, thus such consolidating mapping is applied primarily for VMs.

Due to their energy efficiency benefits, a substantial amount of research is focusing on workload consolidation algorithms. However, the high cost of ownership and management expertise needed for such complex infrastructures, usually it is not possible for researchers to use an actual data centre's infrastructure to test the effectiveness of their algorithms. Therefore, there is a growing need for data centre infrastructure simulators to offer an environment that allows the evaluation of various consolidation algorithms. Such simulators need to model the behaviour and the energy consumption of each actor (e.g. physical hosts, VMs, routers/switches) and of each activity (e.g. VM migration, PM shutdown/startup) involved in the workload consolidation. Amongst these activities, the most researched one is *VM migration*, because it allows moving the state of VMs between PMs, thus it is useful for re-mapping VMs according to the consolidator's needs.

In recent years, several simulators implemented models for VM migration. For example, the work in [2] added a model to the SimGrid [3] simulator that focuses on the migration's performance overhead, but not its energy consumption.

In this paper, we first propose a new energy consumption model to be used in data centre simulators that considers VM migration operations. Our new approach extends on our previous generic VM migration model [4], we implemented it in a Cloud infrastructure simulator called DISSECT-CF [5] which is integrated as a backend of the user-side GroudSim simulator [6]. Compared to other simulators, our new model increases the accuracy of the simulation of VM migration and similar major activities involved in the workload consolidation process. Our ultimate aim is to provide the distributed systems research community with a model that is: (1) easy to implement, and (2) able to capture the behaviour of different types of data centres components.

Our migration model is based on the assumption that the source and target hosts are homogeneous. This assumption mimics the current state of most hypervisors (e.g. Xen, KVM) which prevent VM migration between hosts with incompatible architectures.

We validated our model by comparing it with real-life measurements from various benchmarks executed on VMs migrated across two different sets of hosts in a private Cloud. This way, we managed to: (1) improve the energy models of GroudSim/DISSECT-CF, (2) validate our new model's implementation under different kinds of operational scenarios with and without VM migration, and (3) achieve a 45.2% improvement in normalised error (NRMSE) compared to the state-of-the-art CloudSim simulator.

The paper is organised as follows. First, we review the related work in Section II. We describe in Section III the implementation of our model in the GroudSim and DISSECT-CF simulators, and evaluate its performance and accuracy in Section IV. Finally, we conclude our paper in Section V.

II. Related work

Several papers proposed models for data centre energy consumptions, however, they either focus on a specific CPU architecture or assume a linear relationship between CPU usage and energy consumption.

CloudSim [7] is the most used and cited simulator of different components of the data centre infrastructure, including internal networking and energy consumption. However, CloudSim assumes that the energy consumption exclusively depends on CPU utilization by ignoring other components such as memory and network. Moreover, it does not take into account several important parameters in its VM migration model such as overcommitment and memory dirtying rate.

SimGrid [3] provides a scalable and fast simulation framework of Cloud data centres, Grid and peer-to-peer systems, including a model for simulating VM migration [2]. However, it provides no energy consumption model for VM migration (at the time our paper writing).

GreenCloud [8] offers packet-level simulations for energy-aware Cloud computing data centres. It provides the capability of separately modelling the energy consumption of all data centre components, including CPU, network, and storage. However, its CPU model is based on Xeon processors only and no energy consumption model for live migration is provided.

Our work is based on the GroudSim [6] simulation backend of the ASKALON system [9] that, due to its integration with the DISSECT-CF [10] Cloud infrastructure simulator, provides models for energy consumption of data centre components, as well as for VM migration and networking.

III. Simulation framework

In this section, we describe the simulation framework in which we implement our models consisting of two main parts: GroudSim, which provides the user side of the IaaS Cloud, and DISSECT-CF, which provides the internal infrastructure side.

GroudSim is a Java-based simulation toolkit for scientific applications running on Grid and Cloud infrastructures. GroudSim uses a discrete-event simulation toolkit consisting of a future event list and a time advance algorithm that offers improved performance and scalability compared to other process-based approaches [7]. GroudSim focuses on the user-side of IaaS Cloud computing and is currently used as an additional backend in the ASKALON system enabling users to perform seamless development, debugging, simulation and execution of Grid/Cloud applications using the same interface [9].

GroudSim lacks knowledge of the internal IaaS infrastructure. Since this knowledge is essential for the simulation of energy consumption in data centres, we connected it to DISSECT-CF that is a compact and highly customisable open source Cloud simulator with special focus on the IaaS systems.

A. VM migration model

In our work, we aim to simulate not only power consumption but also the VM migration time. VM migration is the process of transferring the VM state from one *source* host

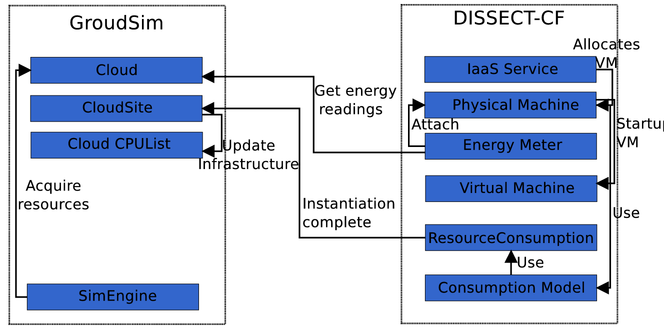

Fig. 1: Interaction between GroudSim and DISSECT-CF.

to another *target* host. We distinguish between two types of VM migration: *non-live* and *live migration*. In the non-live migration, the VM state is transferred after suspending the VM on the source and then resuming it on the target host. In the live migration, the state of the VM is transferred while the VM is still running. For both migration types, we identified in [4] three VM migration phases:

- *Initiation phase*, during which source host prepares transferring the VM state to the target host and the target reserves the resources necessary to host the VM;
- *Transfer phase*, during which the VM state is transferred from the source to the target host in a way depending on whether a non-live or live migration is performed;
- *Activation phase*, during which the source host frees the resources occupied by the VM and the target starts it.

We therefore define the *VM migration time* $T_{\text{migr}}(v, h, \mathcal{S}, \mathcal{T})$ on host h for migrating the VM v from the source $\mathcal{S}$ to the target host $\mathcal{T}$ as the sum of the times required in each phase:

$$T_{\text{migr}}(v, h, \mathcal{S}, \mathcal{T}) = T_{init}(v, h) + T_{\text{transf}}(v, h, \mathcal{S}, \mathcal{T}) + \\ + T_{\text{activ}}(v, h). \quad (1)$$

In the initiation phase, the source host prepares a checkpoint of the VM to be sent to the target. In the activation phase, the source host frees the resource allocated to the VM and the target starts it. Therefore, the times required by both initiation $T_{init}(v, h)$ and activation $T_{activ}(v, h)$ phases are only dependent on the VM size $\texttt{SIZE}(v)$ and the storage bandwidth on the host h:

$$T_{\text{init}}(v, h) \quad = \quad T_{\text{activ}}(v, h) \quad = \quad \frac{\texttt{SIZE}(v)}{\texttt{BW}_{\text{io}}(h)}. \quad (2)$$

The transfer phase, on the other hand, has a different execution time for a live or a non-live migration. The non-live migration time $T_{\text{transf}}^{\text{nonlive}}$ depends only on the VM size and the bandwidth between the two hosts $\texttt{BW}_{\text{net}}(\mathcal{S}, \mathcal{T})$:

$$T_{\text{transf}}^{\text{nonlive}} = \frac{\texttt{SIZE}(v)}{\texttt{BW}_{\text{net}}(\mathcal{S}, \mathcal{T})}. \quad (3)$$

Live migration is instead performed iteratively while the VM v is still running. Therefore, the VM state needs to be continuously updated over a predefined number of iterations, set in the hypervisor's configuration. After the initial state transfer, each iteration transfers only the memory pages that have been modified during the previous transfer of the VM state, leading to the following live VM transfer time:

$$T_{\text{transf}}^{\text{live}} = \frac{\texttt{SIZE}(v)}{\texttt{BW}_{\text{net}}(\mathcal{S}, \mathcal{T})} + \sum_{i=1}^{\mathcal{I}} \frac{\texttt{DP}(v, i)}{\texttt{BW}_{\text{net}}(\mathcal{S}, \mathcal{T})}, \tag{4}$$

where $\mathcal{I}$ is the number of iterations and:

$$\texttt{DP}(v, i) = \frac{\texttt{SIZE}(v)}{\texttt{PS}(v)} \cdot \texttt{DR}(v, i), \tag{5}$$

where $DR(v, i)$ is the dirtying rate of the VM v or the percentage of memory pages marked as dirty during an iteration i, and $PS(v)$ is the size of each memory page of VM v.

Section III gives implementation details of this model.

B. Energy simulation

This section describes the design and implementation of our energy model in the DISSECT-CF infrastructure simulator [5] and integrated in the user-oriented GroudSim simulator [10].

1) Energy modelling: In this module, we extended the `ConsumptionModel` class with two subclasses (`CPUConsumptionModel` and `LinearConsumptionModel`). Each class provides an `evaluateConsumption(double load)` method which, when queried, gives the instantaneous power consumption according to the instantaneous load represented by the `load` parameter that models the relative use of the particular resource (e.g. CPU, network, storage). We calculate the load of CPU resources as follows:

$$load_{\text{cpu}} = \frac{\texttt{CPU}(h, t)}{\texttt{CPU}^{\max}(h)}, \tag{6}$$

where $\texttt{CPU}^{\max}(h)$ is the maximum CPU on the host h and $\texttt{CPU}(h, t)$ is the load of the host h at time instance t. We also define the network load as:

$$load_{\text{net}} = \frac{\texttt{BW}_{\text{net}}(h, t)}{\texttt{BW}_{\text{net}}^{\max}(h)}, \tag{7}$$

where $\texttt{BW}_{\text{net}}^{\max}(h)$ is the maximum bandwidth on the network interface of host h and $\texttt{BW}_{\text{net}}(h, t)$ is the bandwidth on host h at the time instance t. Finally, we define the storage load as:

$$load_{\text{io}} = \frac{\texttt{BW}_{\text{io}}(h, t)}{\texttt{BW}_{\text{io}}^{\max}(h)}, \tag{8}$$

where $\texttt{BW}_{\text{io}}^{\max}(h)$ is the maximum storage bandwidth on the host h and $\texttt{BW}_{\text{io}}(h, t)$ is the bandwidth at the time instance t.

C. GroudSim-DISSECT-CF Interaction

We display in Figure 1 how to obtain DISSECT-CF energy readings in GroudSim. To measure the energy consumption in a data centre, we need two basic information: the physical machines and their load. For this reason, this operation is performed by the *IaaS Service* of DISSECT-CF responsible for both instantiating the data centre infrastructure and allocating a VM to a suitable physical machine. For this purpose, the IaaS Service meter attaches to each host defined in the data centre an `EnergyMeter`. For each host, we define a `ConsumptionModel` for CPU, network and storage that defines the instantaneous power consumption of each component, as discussed in Section III-B1. Energy meters collect these instantaneous power consumption values with a user defined frequency and calculate the energy consumption based on the simulated time spent since the last power measurement. At the end of the simulation, the IaaS service's meter aggregates the energy consumption for the entire data centre.

<table>
<tr><th>Machine</th><th>Available virtual CPUs</th><th>Available RAM</th><th>Gigabit NIC</th><th>Gigabit switch</th></tr>
<tr><td><code>m01</code></td><td rowspan="2">32 (8 Opteron 8356, dual threaded)</td><td rowspan="2">32GB</td><td rowspan="2">Broadcom BCM5704</td><td rowspan="2">Cisco Catalyst 3750</td></tr>
<tr><td><code>m02</code></td></tr>
<tr><td><code>o1</code></td><td rowspan="2">40 (20 Xeon E5-2690, dual threaded)</td><td rowspan="2">128GB</td><td rowspan="2">Intel 82574L</td><td rowspan="2">HP 1810-8G</td></tr>
<tr><td><code>o2</code></td></tr>
</table>

TABLE I: Hardware configuration.

IV. Evaluation

In this section, we evaluate our simulated energy model for VM migrations by first describing the selected benchmarks and the experimental testbed. Then, we describe how we simulate the execution of these benchmarks on the top of DISSECT-CF. Finally, we compare the results of our simulations with energy traces collected from real executions in the simulated Cloud.

A. Benchmarks

For benchmarking our implementation of the model inside our simulator, we employed the benchmarks we designed in [4] to build a new model for VM migration. We made this choice because: (1) they already proved their effectiveness in testing our VM migration model, and (2) they allow us to check the accuracy of our CPU, network and storage models by varying the CPU load and the dirtying rate, which are the parameters that mostly affect the VM migration. Each experiment is executed on both sets of machines described in Table I. On each set of machine, we installed the Xen $4.2.5$ hypervisor. To each machine, we attach an external Voltech PM1000+ power reader. We collected the measurements during the execution of our benchmarks. For each experimental run, the measurements starts after deploying the VMs on the physical machines. The migration is issued once the power consumption of the host stabilises and each measurement ends once the power trace stabilises, meaning that we obtain twenty consecutive power measurements with a difference lower than 0.3%, which is below our measurement device's accuracy.

B. Simulation benchmarking

After collecting the real world traces, we implement the same benchmarks on the top of DISSECT-CF to evaluate the accuracy of our simulations. We configure a micro data centre with two physical machines matching the configuration of the two kinds of machines we used for regression modelling (see in Table I). We simulate the load by deploying VMs on the physical machines and configure each VM to have $4\,\text{GB}$ of memory to resemble the configuration we used to build our traces. To simulate the execution of our benchmarks, we assign

Host	*MAE-NONLIVE*		*NRMSE-NONLIVE*		*MAE-LIVE*		*NRMSE-LIVE*	
	Power [W]	*Energy* [J]	*Power* [%]	*Energy* [%]	*Power* [W]	*Energy* [J]	*Power* [%]	*Energy* [%]
Source (m01)	42.97	4292.9	16.6	16.9	38.45	5345.8	15.5	8.1
Target (m02)	51.97	4179.5	18.3	15.1	67.33	6341.8	22	9.3
Source (o1)	11.98	2248.6	8.2	14.6	27.6	3375.5	18.4	11
Target (o2)	18.19	2417	14.6	13.2	48.1	5518	14.2	25.6

TABLE II: Error for the each machine set.

to each VM computing tasks resembling the execution of the selected workloads.

C. Simulation results

In this section, we compare the results obtained by our simulator with the traces obtained from our experiments. We compute the MAE and NRMSE error metrics on both instantaneous power and energy consumption on both sets of machines. The results are summarised in Table II for each machine set.We observe that our simulation is able to provide power values with a MAE not higher than $67.3\,\mathrm{W}$ compared to the real ones. This value is, however, influenced by the fact that in some cases the power consumption is underestimated by around $100\,\mathrm{W}$ because the test scenarios active during those minutes perform non-live migrations while both hosts are idle In these situations, the simulator only considers the power consumption caused by the network and storage despite some inherent CPU consumption caused by these two operations. Thus, the simulator considers idle CPU consumption for both hosts, despite slight CPU load caused by the need for supporting the storage operations. In future work, we will aim at modelling this inherent CPU load in a generic way to increase the accuracy of the simulator in these unlikely test scenarios too. Nevertheless, NRMSE is in each case between 8% and 22% for instantaneous power consumption, and between 8% and 25% for energy consumption (see Table II showing that our simulator is able to predict both energy and instantaneous power with good accuracy.

V. Conclusion and Future Work

In this paper, we developed a new energy model for data centres energy consumption that considers CPU, network and storage hardware components. We evaluated its accuracy on traces collected from two different types of physical machines in a private Cloud, showing a relative error lower than 18% on both data sets. Afterwards, we implemented our model in the user-side GroudSim simulator by exploiting its integration with DISSECT-CF infrastructure simulator. We evaluated the accuracy of our implementation by comparing it with real measurements, showing a NRMSE between 8% and 22% for power prediction and between 8% and 25.6% for energy estimation. Finally, we compared the results obtained by our implementation with the CloudSim state-of-the-art simulator showing an improvement of at least 36.8% in energy prediction accuracy. In the future, we plan to perform further extensions to our simulator by improving the energy models for network and storage and use them for studying the effects of different energy-aware consolidation algorithms in modern virtualized data centres. We are further interested in validating our simulator with different real-world benchmarks such as TPC-C or SPECPower.

References

[1] L.A. Barroso and U. Holzle. The case for energy-proportional computing. *Computer*, 40(12):33–37, Dec 2007.

[2] T. Hirofuchi, A. Lebre, and L. Pouilloux. Adding a live migration model into simgrid: One more step toward the simulation of infrastructure-as-a-service concerns. In *5th International Conference on Cloud Computing Technology and Science*, volume 1, pages 96–103, 2013.

[3] Henri Casanova, Arnaud Legrand, and Martin Quinson. Simgrid: A generic framework for large-scale distributed experiments. In *Proceedings of the Tenth International Conference on Computer Modeling and Simulation*, UKSIM '08, pages 126–131, Washington, DC, USA, 2008. IEEE Computer Society.

[4] Vincenzo De Maio, Gabor Kecskemeti, and Radu Prodan. A workload-aware energy model for virtual machine migration. In *CLUSTER 2015*, Chicago, Sep 2015. IEEE Computer Society.

[5] Gabor Kecskemeti. DISSECT-CF: a simulator to foster energy-aware scheduling in infrastructure clouds. *Simulation Modelling Practice and Theory*, 58P2:188–218, November 2015.

[6] Simon Ostermann, Kassian Plankensteiner, and Radu Prodan. Using a new event-based simulation framework for investigating resource provisioning in clouds. *Scientific Programming*, 19(2-3):161–178, 2011.

[7] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose, and Rajkumar Buyya. Cloudsim: A toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience*, 41(1):23–50, January 2011.

[8] Dzmitry Kliazovich, Pascal Bouvry, and SameeUllah Khan. Greencloud: a packet-level simulator of energy-aware cloud computing data centers. *The Journal of Supercomputing*, 62(3):1263–1283, 2012.

[9] Simon Ostermann, Kassian Plankensteiner, Radu Prodan, Thomas Fahringer, and Alexandru Iosup. Workflow monitoring and analysis tool for ASKALON. In *Grid and Services Evolution, Proceedings of the 3rd CoreGRID Workshop on Grid Middleware, June 5-6, 2008, Barcelona, Spain*, pages 1–14, 2008.

[10] Gabor Kecskemeti, Simon Ostermann, and Radu Prodan. Fostering energy-awareness in simulations behind scientific workflow management systems. In *Proceedings of the 7th IEEE/ACM International Conference on Utility and Cloud Computing, UCC 2014, London, United Kingdom, December 8-11, 2014*, pages 29–38, 2014.